AMENDMENTS TO THE CLAIMS

The following listing of claims will replace all prior versions and listings of claims in the application.

LISTING OF CLAIMS

1-19. (cancelled)

20. (Currently amended) An information retrieval system for retrieving information a user seeks from a plurality of documents, comprising:

~~document storage means for storing the plurality of documents;~~

feature amount extraction means for extracting a feature amount of each of the plurality of documents <u>stored in document storage means</u>;

clustering means for classifying the plurality of documents into a plurality of clusters based on the extracted feature amounts so that each cluster includes one document or a plurality of documents having feature amounts similar to each other as an element;

cluster term label preparation means for automatically selecting one or more terms, which is or are arranged in order of high term score, as a label of the cluster, for each of the plurality of clusters, the term score being obtained by calculating the number of documents in which a term appears in the cluster, for each of the terms included in documents belonging to the cluster;

document retrieval means for retrieving a document satisfying a retrieval condition input by the user among the plurality of documents; and

interface means for presenting the retrieved document together with the label of the cluster, to which the retrieved document belongs, and the rest of documents belonging to the cluster, as retrieval results,

wherein the feature amount extraction means extract feature vectors as the feature amount,

the feature vector is a vector having as an element a pair of a keyword of each of the plurality of documents stored in the document storage means, and a weight of the keyword, and

the clustering means classify the plurality of documents into a plurality of clusters so that each cluster includes as an element one document among the plurality of documents, in which a ratio between the minimum value and the maximum value of the sum of the weight of the same keywords in the feature amount is large.

21.   (Currently amended) An information retrieval system for retrieving information a user seeks from a plurality of documents, comprising:

~~document storage means for storing the plurality of documents;~~

feature amount extraction means for extracting a feature amount of each of the plurality of documents stored in document storage means;

clustering means for classifying the plurality of documents into a plurality of clusters based on the extracted feature amounts so that each cluster includes one document or a plurality of documents having feature amounts similar to each other as an element;

cluster sentence label preparation means for automatically selecting one sentence as a label of the cluster based on a term score for each of the plurality of clusters, the sentence being included in documents belonging to the cluster, the term score being obtained by calculating the number of documents in which a term appears in the cluster, for each of the terms included in documents belonging to the cluster;

document retrieval means for retrieving a document satisfying a retrieval condition input by the user among the plurality of documents; and

interface means for presenting the retrieved document together with the label of the cluster, to which the retrieved document belongs, and the rest of documents belonging to the cluster, as retrieval results,

wherein the feature amount extraction means extract feature vectors as the feature amount,

the feature vector is a vector having as an element a pair of a keyword of each of the plurality of documents stored in the document storage means, and a weight of the keyword, and

the clustering means classify the plurality of documents into a plurality of clusters so that each cluster includes as an element one document among the plurality of documents, in which a ratio between the minimum value and the maximum value of the sum of the weight of the same keywords in the feature amount is large.

22. (Currently amended) The information retrieval system of Claim 21, wherein, the cluster sentence label preparation means work out a sum of term scores of all terms included in the sentence, and select a sentence in which the sum of the term

scores is largest as a label of the cluster, for each of the sentences included in documents belonging to the cluster, and

when a plurality of sentences in which the sum of the term scores is largest exist, one sentence having the smallest number of characters is selected.

23. (Currently amended) An information retrieval system for retrieving information a user seeks from a plurality of documents, comprising:

~~document storage means for storing the plurality of documents;~~

feature amount extraction means for extracting a feature amount of each of the plurality of documents stored in document storage means;

clustering means for classifying the plurality of documents into a plurality of clusters based on the extracted feature amounts so that each cluster includes one document or a plurality of documents having feature amounts similar to each other as an element;

cluster label preparation means for automatically generating a cluster label representing the contents of the cluster based on terms contained in feature vectors, for each of the plurality of clusters;

document label preparation means for preparing a document label representing the contents of the document, for each of the clustered documents;

document retrieval means for retrieving a document satisfying a retrieval condition input by the user among the plurality of documents; and

interface means for presenting the retrieved document together with the cluster label of the cluster to which the retrieved document belongs, the rest of documents

belonging to the cluster, and the document labels which are associated with each of the retrieved document and the rest of documents, as retrieval results,

wherein the feature amount extraction means extract feature vectors as the feature amount,

the feature vector is a vector having as an element a pair of a keyword of each of the plurality of documents stored in the document storage means, and a weight of the keyword, and

the clustering means classify the plurality of documents into a plurality of clusters so that each cluster includes as an element one document among the plurality of documents, in which a ratio between the minimum value and the maximum value of the sum of the weight of the same keywords in the feature amount is large.

24.   (Canceled)

25.   (Currently amended) The information retrieval system of Claim [[24]] 23, wherein, the document label preparation means selects as the document label one sentence in which the sum of TFIDF values of the terms included in the document is large as the document label based on appearance frequency information of terms among all the sentences included in the document.

26.   (Currently amended) An information retrieval system for retrieving information a user seeks from a plurality of answer documents, comprising:

document storage means for storing the plurality of answer documents and a plurality of question documents, at least one or more question documents being associated with each of the answer documents;

feature amount extraction means for extracting a feature amount of each of the plurality of answer documents;

clustering means for classifying the plurality of answer documents into a plurality of clusters based on the extracted feature amounts so that each cluster includes one document or a plurality of documents having feature amounts similar to each other as an element;

wherein said clustering means is configured to programmatically enforce a rule wherein the number of the plurality of clusters is determined such that the number of clusters having two or more elements of the plurality of clusters is maximized;

question document retrieval means for retrieving a question document conforming with a user question input by the user among the plurality of question documents; and

interface means for presenting the retrieved question document and the answer document associated with the question document together with the rest of answer documents included in the cluster to which the answer document belongs, as retrieval results,

<u>wherein the feature amount extraction means extract feature vectors as the feature amount,</u>

the feature vector is a vector having as an element a pair of a keyword of each of the plurality of documents stored in the document storage means, and a weight of the keyword, and

the clustering means classify the plurality of documents into a plurality of clusters so that each cluster includes as an element one document among the plurality of documents, in which a ratio between the minimum value and the maximum value of the sum of the weight of the same keywords in the feature amount is large.

27.   (Previously presented) The information retrieval system of Claim 26, wherein, the interface means receives selection of an answer document by the user among the answer documents of the presented retrieval results, and

the information retrieval system further comprises document upgrading means for newly storing the document of the user question in the document storage means in association with the selected answer document.

28.   (New) The information retrieval system of Claim 27, wherein, the document updating means newly stores the document of the user question in the document storage means in association with the selected answer document when the similarity of the document of the user question and the question document confirming with the user question is less than a predetermined value.

29.   (New) The information retrieval system of Claim 21, wherein, the cluster sentence label preparation means work out a sum of term scores of all terms included in

the sentence, and select a sentence in which the sum of the term scores is largest as a label of the cluster, for each of the sentences included in documents belonging to the cluster, and

when a plurality of sentences in which the sum of the term scores is largest exist, a sentence the head of which is located nearest to the beginning of the document is selected.